

WEST Search History

DATE: Tuesday, July 15, 2003

Set Name Query

side by side

Hit Count Set Name
result set

DB=USPT; THES=ASSIGNEE; PLUR=YES; OP=OR

L14	L8 and catalog\$	4	L14
L13	L7 and ((determin\$ or select\$ or pick\$ or choos\$) same (amount or quantit\$))	0	L13
L12	L7 and ((determin\$ or select\$ or pick\$ or choos\$) with (amount or quantit\$))	0	L12
L11	L9 and l5	1	L11
L10	L9 and l7	0	L10
L9	((705/26)!.CCLS.)	797	L9
L8	L7 and quantit\$	4	L8
L7	L5 and ((history or past\$ or profil\$) same (command or quer\$))	4	L7
L6	L5 and ((history or past\$ or profil\$) with quer\$)	0	L6
L5	L4 and ((display\$ or review\$) with log\$)	12	L5
L4	L3 and @ad<=20000630	35	L4
L3	(automated adj patent adj system) and aps and search\$	36	L3
L2	L1 and @ad<=20000630	3709	L2
L1	(automated patent system) and aps and search\$	4012	L1

END OF SEARCH HISTORY

End of Result Set



Generate Collection

Print

L26: Entry 1 of 1

File: USPT

Oct 1, 2002

US-PAT-NO: 6460034

DOCUMENT-IDENTIFIER: US 6460034 B1

TITLE: Document knowledge base research and retrieval system

DATE-ISSUED: October 1, 2002

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Wical; Kelly	San Carlos	CA		

ASSIGNEE-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY	TYPE CODE
Oracle Corporation	Redwood Shores	CA			02

APPL-NO: 08/ 861961 [PALM]

DATE FILED: May 21, 1997

INT-CL: [07] G06 F 17/30

US-CL-ISSUED: 707/5

US-CL-CURRENT: 707/5

FIELD-OF-SEARCH: 707/1-5

PRIOR-ART-DISCLOSED:

U.S. PATENT DOCUMENTS

Search Selected

Search ALL

	PAT-NO	ISSU DATE	PATENTEE-NAME	US-CL
<input type="checkbox"/>	5159667	October 1992	Borrey et al.	395/148
<input type="checkbox"/>	5167011	November 1992	Priest	395/54
<input type="checkbox"/>	5257185	October 1993	Farley et al.	364/419.19
<input type="checkbox"/>	5276616	January 1994	Kuga et al.	364/419.08
<input type="checkbox"/>	5325298	June 1994	Gallant	364/419.19
<input type="checkbox"/>	5369763	November 1994	Biles	395/600
<input type="checkbox"/>	5418948	May 1995	Turtle	707/4
<input type="checkbox"/>	5418961	May 1995	Segal et al.	707/1
<input type="checkbox"/>	5442780	August 1995	Takanashi et al.	395/600
<input type="checkbox"/>	5488725	January 1996	Turtle et al.	707/5
<input type="checkbox"/>	5544049	August 1996	Henderson et al.	704/7
<input type="checkbox"/>	5598557	January 1997	Doner et al.	707/5
<input type="checkbox"/>	5619709	April 1997	Caid et al.	707/532
<input type="checkbox"/>	5625767	April 1997	Bartell et al.	395/140
<input type="checkbox"/>	5630117	May 1997	Oren et al.	395/602
<input type="checkbox"/>	5630125	May 1997	Zellweger	395/614
<input type="checkbox"/>	5634051	May 1997	Thomson	707/5
<input type="checkbox"/>	5642502	June 1997	Driscoll	707/5
<input type="checkbox"/>	5675819	October 1997	Schuetze	704/10
<input type="checkbox"/>	5708829	January 1998	Kadashevich et al.	707/5
<input type="checkbox"/>	5717914	February 1998	Husick et al.	707/5
<input type="checkbox"/>	5721902	February 1998	Schultz	707/4
<input type="checkbox"/>	5724567	March 1998	Rose et al.	707/2
<input type="checkbox"/>	5761418	June 1998	Francis et al.	395/200.31
<input type="checkbox"/>	5768578	June 1998	Kirk et al.	707/100
<input type="checkbox"/>	5787234	July 1998	Molloy	706/46
<input type="checkbox"/>	5787417	July 1998	Hargrove	707/4
<input type="checkbox"/>	5819258	October 1998	Vaithyanathan et al.	707/2
<input type="checkbox"/>	5842206	November 1998	Sotomayor	707/5
<input type="checkbox"/>	5875446	February 1999	Brown et al.	707/3
<input type="checkbox"/>	5956708	September 1999	Dyko et al.	707/3
<input type="checkbox"/>	5963940	October 1999	Liddy et al.	707/5
<input type="checkbox"/>	5978799	November 1999	Hirsch	707/4
<input type="checkbox"/>	6006221	December 1999	Liddy et al.	707/5
<input type="checkbox"/>	6026388	February 2000	Liddy et al.	707/1

OTHER PUBLICATIONS

Liddy, Elizabeth D. et al., "DR-LINK System: Phase I Summary" Proceedings of the

TIPSTER, Syracuse University, NY, Sep. 1993, pp. 93-110, 114.*
Liddy, Elizabeth D. et al., "DR-LINK: A System Update for TREC-2," TREC Text Retrieval Conf, NIST, Wash., D.C., Aug. 1993, pp. 85-99.*
S. Al-Hawamdeh, "Compound Document Processing," Proc. of the 15th Annual Internat'l Computer Software & Applications Conf., Sep. 1991, p. 640-4.*
Cox, John "'Text-Analysis' Server to Simplify Queries", Communications Week, Apr. 19, 1993.

ART-UNIT: 2171

PRIMARY-EXAMINER: Von Buhr; Maria N.

ATTY-AGENT-FIRM: Stattler Johansen & Adeli LLP

ABSTRACT:

A knowledge base search and retrieval system, which includes factual knowledge base queries and concept knowledge base queries, is disclosed. A knowledge base stores associations among terminology/categories that have a lexical, semantical or usage association. Document theme vectors identify the content of documents through themes as well as through classification of the documents in categories that reflects what the documents are primarily about. The factual knowledge base queries identify in response to an input query, documents relevant to the input query through expansion of the query terms as well as through expansion of themes. The concept knowledge base query does not identify specific documents in response to a query, but specifies terminology that identifies the potential existence of documents in a particular area.

25 Claims, 22 Drawing figures

End of Result Set



Generate Collection

Print

L26: Entry 1 of 1

File: USPT

Oct 1, 2002

DOCUMENT-IDENTIFIER: US 6460034 B1

TITLE: Document knowledge base research and retrieval system

Abstract Text (1):

A knowledge base search and retrieval system, which includes factual knowledge base queries and concept knowledge base queries, is disclosed. A knowledge base stores associations among terminology/categories that have a lexical, semantical or usage association. Document theme vectors identify the content of documents through themes as well as through classification of the documents in categories that reflects what the documents are primarily about. The factual knowledge base queries identify, in response to an input query, documents relevant to the input query through expansion of the query terms as well as through expansion of themes. The concept knowledge base query does not identify specific documents in response to a query, but specifies terminology that identifies the potential existence of documents in a particular area.

US Patent No. (1):

6460034

Brief Summary Text (6):

In response to a query, a word match based search and retrieval system parses the repository of information to locate a match by comparing the words of the query to words of documents in the repository. If there is an exact word match between the query and words of one or more documents, then the search and retrieval system identifies those documents. These types of prior art search and retrieval systems are thus extremely sensitive to the words selected for the query.

Brief Summary Text (7):

The terminology used in a query reflects each individual user's view of the topic for which information is sought. Thus, different users may select different query terms to search for the same information. For example, to locate information about financial securities, a first user may compose the query "stocks and bonds", and a second user may compose the query "equity and debt." For these two different queries, a word match based search and retrieval system would identify two different sets of documents (i.e., the first query would return all documents that have the words stocks and bonds and the second query would return all documents that contain the words equity and debt). Although both of these query terms seek to locate the same information, with a word search and retrieval system, different terms in the query generate different responses. Thus, the contents of the query, and subsequently the response from word based search and retrieval systems, is highly dependent upon how the user expresses the query term. Consequently, it is desirable to construct a search and retrieval system that is not highly dependent upon the exact words chosen for the query, but one that generates a similar response for different queries that have similar meanings.

Brief Summary Text (8):

Prior art search and retrieval systems do not draw inferences about the true content of documents available. If the search and retrieval system merely compares words in a document with words in a query, then the content of a document is not really being compared with the subject matter identified by the query term. For example, a restaurant review article may include words such as food quality, food presentation, service, etc., without expressly using the word restaurant because the topic, restaurant, may be inferred from the context of the article (e.g., the restaurant review article appeared in the dining section of a newspaper or travel magazine).

For this example, a word comparison between a query term "restaurant" and the restaurant review article may not generate a match. Although the main topic of the restaurant review article is "restaurant", the article would not be identified. Accordingly, it is desirable to infer topics from documents in a search and retrieval system in order to truly compare the content of documents with a query term.

Brief Summary Text (12):

Factual or document knowledge base query processing in a search and retrieval system identifies, in response to a query, a plurality of documents relevant to the query. The search and retrieval system stores content information that defines the overall content for a repository of available documents. To process a query, which includes at least one query term, the factual knowledge base query processing selects documents relevant to the query terms. In addition, the factual knowledge base query processing selects additional documents, through use of the content information, such that the additional documents have content in common with content in the original document set.

Brief Summary Text (15):

In one embodiment, the knowledge base includes a plurality categories. The search and retrieval system stores document theme vectors that classify the documents, as well as the themes identified for the documents, in categories of the knowledge base. The factual knowledge base query processing maps the expanded query term set to categories of the knowledge base, and selects documents, as well as the document themes, classified for those categories. Additional documents, which have themes common to the themes of the original document set, are selected. From all of the themes selected, theme groups, which include themes common to more than one document, are generated. Documents from theme groups are identified in response to the query. Also, the theme groups are ranked based on the relevance of the theme groups to the query. Furthermore, documents, within the theme groups, are ranked in order of relevance to the query.

Detailed Description Text (3):

The search and retrieval system of the present invention utilizes a rich and comprehensive content processing system to accurately identify themes that define the content of the source material (e.g., documents). In response to a search query, the search and retrieval system identifies themes, and the documents classified for those themes. In addition, the search and retrieval system of the present invention draws inferences from the themes extracted from a document. For example, a document about wine, appearing in a wine club magazine, may include the words "vineyards", "Chardonnay", "barrel fermented", and "french oak", which are all words associated with wine. As described more fully below, if the article includes many content carrying words that relate to the making of wine, then the search and retrieval system infers that the main topic of the document is about wine, even though the word "wine" may only appear a few times, if at all, in the article. Consequently, by inferring topics from terminology of a document, and thereby identifying the content of a document, the search and retrieval system locates documents with the content that truly reflect the information sought by the user. In addition, the inferences of the search and retrieval system provide the user with a global view of the information sought by identifying topics related to the search query although not directly included in the search query.

Detailed Description Text (5):

As described more fully below, the search and retrieval system of the present invention maps search queries to all senses, and presents the results of the query to reflect the contextual mapping of the query to all possible senses. In one embodiment, the search and retrieval system presents the results relative to a classification system to reflect a context associated with the query result. For example, if the user search term is "stock", the search and retrieval system response may include a first list of documents under the category "financial securities", a second list of documents under the category "animals", and a third category under the category "race automobiles." In addition, the search and retrieval system groups categories identified in response to a query. The grouping of categories further reflects a context for the search results. Accordingly, with contextual mapping of the present invention, a user is presented with different contextual associations in response to input of a search query that has more than one sense.

Detailed Description Text (6):

In one embodiment, the search and retrieval system of the present invention includes factual knowledge base queries as well as concept knowledge base queries. As described more fully below, the factual knowledge base queries identify, in response to a query, the relevant themes, and the documents classified for those themes. In contrast, the concept knowledge base queries do not identify specific documents in response to a query, but identify the potential existence of a document by displaying associated categories and themes. In essence, for the concept knowledge base query, the user learns that if documents do exist for the search query, the documents may be located under or associated with these categories and terminology for a particular context of the search query. The user may use the identified categories and terminology to locate the information in a different system. For example, the search and retrieval system may operate on a repository that includes a limited set of documents. If a user is unsuccessful in locating a document in this repository, then the categories and associated terminology learned from a concept query may be used to search a larger repository, such as the Internet.

Detailed Description Text (9):

FIG. 1 is a block diagram illustrating one embodiment for the search and retrieval system of the present invention. In general, the search and retrieval system 100 receives, as input, user queries, and generates, as output, search results which, depending upon the mode of operation, identifies categories and documents. The search and retrieval system 100 is cataloged with one or more documents, labeled documents 130 on FIG. 1. The documents 130 may include a compilation of information from any source. For example, the documents 130 may be information stored on a computer system as computer readable text. Also, the documents 130 may be accessed via a network, and stored at one or more remote locations. The content of the documents 130 may include articles, books, periodicals, etc.

Detailed Description Text (30):

FIG. 3 illustrates a response to an example query configured in accordance with one embodiment of the search and retrieval system. The example query, shown in block 610 in FIG. 3, includes the query terms "Legal", "Betting", and "China." With this query, the user wants to learn about the legal aspects of betting in China. In response to the query, the search and retrieval system 100, utilizing the knowledge base 155, identifies terminology related to the query terms. Specifically, for this example, the search and retrieval system 100 identifies terminology shown in blocks 620, 630, and 645 of FIG. 3. In general, the knowledge base 155 is used to identify terminology that has a lexical, semantic, or usage association with the query terms. For this example, the knowledge base 155 indicates that the query term "legal" has a sense association with the terms "government", "patents", and "crime." For the query term betting, the knowledge base 155 identifies the terms "casino", "slot machines", and "wagering." Furthermore, a sense association is identified between the terms "Asia" and "Japan" and the query term "China."

Detailed Description Text (31):

As shown in the example of FIG. 3, the search and retrieval system groups terms related to the query based on information identified. For this example, the search and retrieval system 100 identified information in two documents that included the topics "government", "casino", and "Asia." For the second entry, the search and retrieval system 100 identified six documents that included topics on "patents", "slot machines", and "Japan." Furthermore, there were three documents that included information on the topics "crime", "wagering", and "China."

Detailed Description Text (33):

The response to a multiple term query may also identify groupings that relate to less than all of the query terms. For example, if the two documents related to the terms "government", "casino", "Asia" did not include a theme classified, or subclassified, under the category Asia, then a grouping of the terms "government and casino" would be displayed. Accordingly, the search and retrieval system presents the most relevant groupings of terms that include information contained in the documents.

Detailed Description Text (34):

For the embodiment shown in FIG. 3, the documents, which include information on the corresponding terms, are presented relative to a classification system. For the grouping of terms "government, casino, Asia", encompassed in block 620, two documents were classified in the category "gaming industry", as shown in block 625. Although the two documents contain information, or themes, about the terms "government, casino, Asia", these two documents were classified in the category

"gaming industry", thereby indicating that the two documents are primarily about the gaming industry. The categories displayed in response to a query permit a user to view the main or central topic of the documents having material on the corresponding terms. As shown in blocks 635 and 640, respectively, four documents were classified under the category "patent law", and two documents were classified under the category "gaming industry." As shown in blocks 650 and 655, one document was classified in the category "insects", and two documents were classified under the category "conservation and ecology." Accordingly, the display of categories in the response to a query alerts the user to the general or most important topic of the documents identified.

Detailed Description Text (35):

The example presentation shown in FIG. 3 provides a global view of the response to the users query. The terms, associated with the query terms, (e.g., blocks 620, 630, and 645), allow the user to visualize groupings of topics located in response to the query. As shown in block 630 of FIG. 3, the search and retrieval system located documents that cover information about patents, slot machines, and Japan. At a simple glance, the user may determine whether the documents located are about the general subject matter for which the user seeks to locate information. For this example, the user may seek information about illegal betting on insects in China. Although the documents that include information on patents, slot machines and Japan is relevant to the search query, (e.g., legal, betting, China), the user is immediately alerted that these documents do not contain material on illegal betting on insects in China. Instead, the terms "crime, wagering, China" immediately steer the user to documents about illegal wagering in China. In addition, the presentation of documents relative to categories of a classification system permits the user to identify the document that is primarily about insects. Thus, because the user has a global view of the query response, the user may immediately identify documents most pertinent to the area for which information is sought. For this example, if the user seeks information on illegal betting in China, then the user is directly steered to the document classified under the category "Insects" and under the topic group of "Crime", "Wagering", and "China."

Detailed Description Text (36):

Accordingly, the search and retrieval system 100 first displays a global view of the response to the query. Thus, the user may focus on the area for which the user seeks information by obtaining a specific context for the query from the presentation of the topic groupings (e.g., 620, 630 and 645). With this response presentation, the user is not required to read a document identified in the query response to determine the overall relevance of the document to the specific information sought by the user. Once the user identifies specific information of interest, the user selects one or more documents by selecting (e.g., double clicking with a cursor control device) the category (e.g., 625, 635, 640, 650, and 655) of interest. For the above example, if the user sought information concerning the illegal betting on insects in China, then the user would select the single document classified under the category "Insects", labeled 650 on FIG. 3, to view the document contents.

Detailed Description Text (47):

In one embodiment, query terms or query phrases are processed to identify the thematic content of terms of the input queries. In general, query term processing involves analyzing the query phrase or terms to determine the most important thematic information in the query terms. In one embodiment, the query processing assigns or generates a query strength to each term, wherein the query strength indicates the relative thematic importance among terms or words in the query. For example, a user may input to the search and retrieval system 100 the phrase "pollution caused by European space stations." For this example, the query processing analyzes the input query to determine that the terms "pollution" and "space stations" are the most important, followed by the term "Europe." The term "cause" receives a much lower query term strength, and the word "by" is completely eliminated for purposes of analysis.

Detailed Description Text (50):

The knowledge base is used to expand the query terms to identify an expanded set of query terms as shown in block 405 of FIG. 5. In general, the query terms are mapped to categories in the knowledge base. The directed graph of the knowledge base is then used to identify relevant categories/terms to expand the query term set to include related categories/terms. FIG. 6 illustrates one embodiment for expanding query terms using the knowledge base. Specifically, FIG. 6 shows a portion of a generalized directed graph that includes a plurality of categories/terms with

related categories/ms.

Detailed Description Text (60):

As shown in block 455 of FIG. 5, the process selects the top themes based on a predetermined criteria. For one embodiment, the process selects themes based on a predetermined number of themes or based on a minimum total theme strength. For this example, the factual knowledge base query processing selects only themes identified for more than one document. Thus, "chateaus" is eliminated.

Detailed Description Text (69):

As shown in block 520, the query terms are mapped to the knowledge base (i.e., the query terms are matched to category/terminology in the knowledge base). The query terms are then expanded through use of the knowledge base as shown in block 530 and as described above in conjunction with FIG. 6. As shown in block 540, theme sets associated with the category/terminology for the expanded query terms are selected. Through use of the knowledge base, the selected themes are expanded through a technique similar to the one used to expand the query terms as shown in block 550. As shown in block 560, sets of common denominators of expanded themes are identified for the expanded query terms. For example, if a query includes three terms, then common denominator themes that satisfy all three query terms or their expanded set, are identified. The groups of query terms, groups of expanded query terms, and the corresponding themes are relevance ranked based upon a predetermined criteria as shown in block 570. As shown in block 580, the response to the query is displayed to show: the query terms entered by the user; the groups of query terms selected from the expanded query term set; and the themes organized under the groups of expanded query terms selected.

Detailed Description Text (89):

The search and retrieval system selects common denominators of expanded themes among the expanded query terms to satisfy as many parts of the input query as possible. (FIG. 7, block 560). The search and retrieval system compares themes among the different query terms to identify common denominators. For this example, themes identified for the query term "foods" (Table 12), themes for the query term "Western Europe" (Table 13), and themes for the query term "festivals" (Table 11) are compared. From this comparison, theme groups are extracted.

Detailed Description Text (90):

FIG. 9c illustrates one embodiment for a search and retrieval response in accordance with the example query input. As shown in FIG. 9c, three groups, which satisfy at least a portion of the input query, were identified. Group IA and IB all included themes found in expanded terms for the "festivals", "food", and "Western Europe" input query terms. Specifically, for group IA, the themes: beer, knockwurst, Oktoberfest, stein and sauerkraut, all appear under the categories "customs and practices", "drinking and dining", and "Germany." The expanded query term "customs and practices" maps to "festivals", the "drinking and dining" expanded query term maps to the "foods" category, and the extended query term category "Germany" maps to the query term "Western Europe." For group IB, the extended query terms "festivals", "drinking and dining", and "France" include the themes Mardi Gras, crepes, Camembert, croissant, brie, tripe sausage, onion soup and chicken cordon bleu. Thus, similar to group IA, all distinct parts of the query were satisfied with the corresponding list of themes. For group IIA, two query terms were satisfied (e.g., festivals and foods). The extended query categories "ancient Rome" and "wines" both contain themes for wine, grapes, fermentation, barrels and vineyards.

Detailed Description Text (95):

For the embodiment shown in FIG. 9c, the response to a query includes three types of information. First, identified by the roman numeral headings, the distinct portions of the input query that were satisfied are displayed. For example, for II, the search and retrieval system displays "festivals" and "foods." The categories followed by the capital letters identify to the user the categories used to satisfy the input query. For the example group IA, although the user input the category "foods", the search and retrieval system responded using the category "drinking and dining."

Detailed Description Text (96):

As shown by the above example, the concept knowledge base query does not identify specific documents available in the search and retrieval system 100. Instead, only areas for which potential documents may be classified are identified to the user. A user may then select documents from the categories identified in response to the

concept knowledge base query. If documents in the identified areas are not available, then the user may search other resources, such as the Internet, through use of the categories/topics identified in the response to the concept knowledge base query. Thus, the concept knowledge base query provides a map to the user of potential areas for which a query response may locate information relevant to the query.

Detailed Description Text (98):

FIGS. 10a and 10b illustrate example display responses for the search and retrieval system to the search query "Internet." In response to the Internet query, the search and retrieval system located fifteen documents classified for the category "computer networking." Also, the search and retrieval system identified the terms "Internet Credit Bureau, Incorporated", "Internet Fax Server", "Internet Productions, Incorporated", and "Internet Newbies." As discussed above, for a concept knowledge base query, the results are based on the query mapped to the knowledge base 155. Although no documents were classified under the terms "Internet Credit Bureau, Incorporated", "Internet Fax Server", "Internet Productions, Incorporated", and "Internet Newbies", the terms relate to the search query. The terms are displayed based on the relevance to the search term "Internet." For this embodiment, the relevancy system, indicated by the number of stars, indicates that the category "computer networking" is the most relevant to the query term "Internet."

Detailed Description Text (110):

The profile inquiry has application for use on the Internet. The Internet permits access to a vast amount of information; however, a user is required to sift through the information without any particular guidance as to the context for which a subject appears. With use of the profile query of the present invention, the user receives the context of information available so that the user may readily identify the most relevant information.

CLAIMS:

1. A method for processing queries in a search and retrieval system, said method comprising the steps of: storing a plurality of themes for a repository of documents, wherein each theme for a document defines subject matter disclosed in a document, such that said themes stored for a document define the overall content for said document; processing a query, which includes at least one query term, to select at least one document relevant to said at least one query term; identifying said themes stored for said at least one document selected; and selecting, in response to said query, at least one additional document, not previously selected, that comprises at least one theme in common with said themes identified in said documents selected.

13. A computer readable medium comprising a plurality of instructions, which when executed by a computer, causes the computer to perform the steps of: storing a plurality of themes for a repository of documents, wherein each theme for a document defines subject matter disclosed in a document, such that said themes stored for a document define the overall content for said document; processing a query, which includes at least one query term, to select at least one document relevant to said at least one query term; identifying said themes stored for said at least one document selected; and selecting, in response to said query, at least one additional document, not previously selected, that comprises at least one theme in common with said themes identified in said documents selected.

25. A computer system comprising: memory for storing a plurality of themes for a repository of documents, wherein each theme for a document defines subject matter disclosed in a document, such that said themes stored for a document define the overall content for said document; and a processor unit for processing a query, which includes at least one query term, to select at least one document relevant to said at least one query term, to identify said themes for said at least one document selected, and to select, in response to said query, at least one additional document, not previously selected, that comprises at least one theme in common with said themes identified in said documents selected.